〈Research Paper〉

Journal of the Korean Insitute of Industrial Engineers
Vol. 1 No. 1 June 1975

The Generation of Poisson Random Variates

Chae Ha Pak*

Abstract

Three approximation methods for generating outcomes on Poisson random variables are discussed. A comparison is made to determine which method requires the least computer execution time and to determine which is the most robust approximation. Results of the comparison study suggest the method to choose for the generating procedure depends on the mean value of Poisson random variable which is being generated.

Ⅰ. Introduction

It is frequently desired to generate Poisson random variates in simulations. There are standard exact methods for doing this; the problem arises when a computer is used to generate the Poisson random number which has a large mean. For example, generating one random number such as 105 from a Poisson distribution with mean 100 needs at least 105 calls to a pseudo-random number (uniform(0, 1)) generator. Computer time requirements become important cost factors when considering various methods for generating random numbers.

The objective of this paper is to examine several approximated ways of generating Poisson random variates and to determine the method which gives minimum execution time and small mean squared deviation according to the Poisson mean value. The mean value is the only parameter in the distribution. Comparison statistics to determine the best approximation to the Poisson distribution are the cumulative probability, mean squared deviation and Kolmogorov-Smirnov test. Finally a ccmposite generating procedure according to the mean value is suggested.

The following notation is used in this study

U_i denotes a uniform(0, 1) random variable

N denotes a Poisson random variable where mean is m;

Z denotes a random variable from a standard normal distribution.

Ⅱ. Generation of Poisson- Distributed Variates

A. the Poisson Distribution

A random variable N with integer values has a Poisson distribution if

$$\text{Prob } \{N=n\}=\frac{e^{-m}m^n}{n!} \quad n=0,1,2\ldots$$

In order to generate a Poisson random number N from a Poisson distribution with mean m, the following algorithm is presented. It is the standard exact method for generating these variates.

Let $U_i, i=1,2,\cdots,$ be independent uniform (0, 1) random variate. The Poisson variate, N, is computed as:

$$N=\begin{cases}0 & \text{if } U_i\leq e^{-m};\ i=1,2,3,\cdots \\ n & \text{if } \prod_{i=1}^{n} U_i>e^{-m}\geq\prod_{i=1}^{n+1} U_i\end{cases}$$

* Systems Analysis Group, the Republic of Korea Navy.

where N is distributed as a Poisson with mean m, i.e., Prob $\{N=n\}=e^{-m}m^n/n!$.

Equivalently, since the logarithm is a monotone transformation, we have

$$N=\begin{cases}0 & \text{if } \ln(U_i)\le -m,\\ n & \text{if } \sum_{i=1}^{n}\ln(U_i)>-m\ge\sum_{i=1}^{n+1}\ln(U_i)\end{cases}$$

Letting $E_i=-\ln(U_i)$

$$N=\begin{cases}0 & \text{if } E_i\ge m\\ n & \text{if } \sum_{i=1}^{n}E_i<m\le\sum_{i=1}^{n+1}E_i\end{cases}$$

where the E_i is exponentially distributed variate with mean 1.

If n multiplications of uniform(0, 1) random numbers is strictly greater than e^{-m} and if $n+1$ multiplications of uniform(0, 1) random number is equal to or less than e^{-m} then n is the Poisson random number. Generally, generating one random number from Poisson process with parameter m requires on the average $m+1$ uniform(0, 1) random numbers. This is because the number generated is $n+1$. When m is large it is clear that generating Poisson random numbers with the above method, although it is exact, takes a lot of computer time and this method may be uneconomical. In addition, the large number of multiplications can produce serious precision problems on a digital computer.

B. Approximations for Poisson Variates

1. Normal Approximation

In a Poisson process with parameter λ it is necessary to generate random variables from the Poisson distribution with parameter(mean) m. Now look at counts in $(0, x)$, where x satisfies $\lambda x=m$. The central limit theorem says that as m goes to positive infinity, (or when x goes to the infinity in Poisson process with fixed λ), then N, which has mean m and variance m, is such that $(N+0.5-m)/m^{1/2}$ is approximately distributed as a standard normal random variable. Denote a random variable from a unit normal distribution by Z. So N is distributed approximately as $m^{1/2}Z+m-0.5$ In order to generate Poisson random numbers from the normal distribution, first generate Z; then let

$$N=\begin{cases}0 & \text{if } m^{1/2}Z+m-0.5<1\\ [m^{1/2}Z+m-0.5] & \text{otherwise}\end{cases}$$

where [a] denotes the greatest integer less than or equal to a. N is then the approximated Poisson random variable.

2. Square Root Transformation of Poisson Distribution

If N is a Poisson random variable with mean m then $Y=\sqrt{N+3/8}$ is approximately distributed as a normal distribution with mean $m^{1/2}$ and variance 1/4. This result is due to Bartlett [9].

This method is derived as follows: let $Y=\sqrt{N+C}\sim N(\mu,\sigma^2)$ where C is a non-negative constant. Let $t=N-m$ and $m'=m+C$. Define coefficients for $s=1, 2, 3, \cdots$ by

$$A_s=(-1)^{s+1}\frac{1\cdot(-1)\cdot(-3)\cdots(-2s+3)}{2^s\cdot s!}$$

Then for any $t\ge -m'$ we have a Taylor series expansion.

$$Y=\sqrt{m'}\left\{1+A_1\frac{t}{m'}-A_2\left(\frac{t}{m'}\right)^2+\cdots+(-1)^sA_{s-1}\left(\frac{t}{m'}\right)^{s-1}\right\}+R_s$$

If $t>0$, we see at once $|R_s|<A_st^s/(m')^{s-1/2}$ converges and is bounded [4]. We note now that the moments of t are $\mu_1=0$, $\mu_2=m$, $\mu_3=m$, $\mu_4=3m_2+m$, $\cdots$, which give

$$\text{Var }(Y)\sim\frac{1}{4}\left(1+\frac{3-8C}{8m}+\frac{32C^2-52C+17}{32m^2}\right),$$

so that when $C=3/8$, $\text{Var}(Y)\sim(1+1/16m^2)/4$. Also

$$E(Y)\sim\sqrt{m+C}-\frac{1}{8m^{1/2}}+\frac{24C-7}{128m^{3/2}}$$

Let XNR$=\sqrt{N+3/8}$. Then XNR is approximately normally distributed with mean $\sqrt{m}$ and variance 1/4.

$$Z=\frac{\text{XNR}-\sqrt{m+3/8}}{1/2},$$

$$\text{XNR}=\frac{Z}{2}+\sqrt{m+3/8},$$

$$\sqrt{N+3/8}=\frac{Z}{2}+\sqrt{m+3/8},$$

thus set

$$N=\begin{cases}0 & \text{if}\left(\frac{Z}{2}+\sqrt{m+3/8}\right)^2-3/8<1\\ \left[\left(\frac{Z}{2}+\sqrt{m+3/8}\right)^2-3/8\right] & \text{otherwise}\end{cases}$$

N is then the approximated Poisson variate.

We now need to calculate the probability distribution of $\tilde{N}$ obtained in this way from the square root transformation. We want the probability that

$$n-1+3/8<(\text{XNR})^2<n+3/8$$

if we divide by the variance 1/4 we get

$$4(n-1+3/8)<4(\text{XNR})^2<4(n+3/8)$$

Note that $4\cdot(\text{XNR})^2$ is distributed as a non-central $\chi^2{}_1$ distribution with 1 degree of freedom. The non-central χ^2 density is

$$f_X(x)=\frac{e^{-1/2(x+\mu^2/\sigma^2)}}{2\sqrt{x}\ \sqrt{2\pi}}\left[e^{-\frac{\mu}{\sigma}\sqrt{x}}+e^{\frac{\mu}{\sigma}\sqrt{x}}\right]$$

Thus if $\mu=m^{1/2}$, $\sigma^2=(1/2)^2$, then

$$f_X(x)=\frac{e^{-1/2(x+m/1/4)}}{2\sqrt{x}\ \sqrt{2\pi}}[e^{-(\sqrt{m}/1/2)\sqrt{x}}+e^{(\sqrt{m}/1/2)\sqrt{x}}]$$

and

$$\text{Prob}(X=n)\approx\int_{\frac{n+3/8-1}{(1/2)^2}}^{\frac{n+3/8}{(1/2)^2}}f_X(x)dx$$

This allows us to evaluate directly how well the distribution of $\tilde{N}$ approximates the distribution of a Poisson variate with parameter m.

Note that since in the LLRANDOM package it takes the same amount of time to generate 5 uniform random variables as it takes to generate a normal random variable, the procedures will be competitive timewise once m is much greater than 5.

3. Cube Root Transformation of Poisson Distribution

If N is a Poisson random variable with mean m then $Y=\sqrt[3]{N-1/24}$ is approximately distributed as a normal distribution with mean $m^{1/3}$ and variance $1/9\sqrt[3]{m}$ when $N\geq1$. This comes from the following derivation which is essentially the same procedure as for the square root transformation. Suppose $Y=\sqrt[3]{N+C}$ is distributed as $N(\mu,\sigma^2)$. Let $t=N-m$ and $m'=m+C$. Define coefficients for $s=1,2,3,\cdots$ by

$$a_s=\frac{(-1)^{s+1}1\cdot(-2)(-5)(-8)\cdots(4-3s)}{3^s\cdot s!}$$

For any $t\geq-m'$ we have the Taylor series expansion

$$Y=\sqrt[3]{m'}\left\{1+a_1\frac{t}{m'}-a_2\left(\frac{t}{m'}\right)^2+a_3\left(\frac{t}{m'}\right)^3-a_4\left(\frac{t}{m'}\right)^4\cdots+(-1)^s\ a_{s-1}\left(\frac{t}{m'}\right)^{s-1}\right\}+R_s$$

if $t>0$, we see at once that $|R_s|<a_st_s/(m')^{s-1/3}$ converges. Therefore,

$$R_s=\frac{f^{(s)}\left(1+\theta\frac{1}{m'}\right)}{s!}\left(\frac{t}{m}\right)^s$$

$$0\leq\theta\leq1,\ |t|\leq m'$$

$$R_s(m')^{-1/3}=\left(1+\frac{t}{m'}\right)^{1/3}-\left\{1+a_1\frac{t}{m'}-\cdots+(-1)^s\ a_{s-1}\left(\frac{t}{m}\right)^{s-1}\right\}$$

$$\frac{R_s(m')^{-1/3}}{t^s}=\sum_{i=s}^{\infty}(-1)^{i+1}\ a_i\left(\frac{t}{m'}{}^i\right)$$

converges and is bounded.

We note now that the moments of t are $\mu_1=0$, $\mu_2=m$, $\mu_3=m$, $\mu_4=3m+m,\cdots$, giving us

$$E(Y)\cong\sqrt[3]{m+C}-\frac{1}{18}\cdot\frac{1}{m^{2/3}}+\cdots$$

and

$$\text{Var}(Y)\cong(\sqrt[3]{m'})^2\left\{\frac{1}{9}\cdot\frac{1}{(m')^2}xm-\frac{1}{81}\cdot\frac{2m^2+m}{(m')^4}\cdots\right\}=\frac{1}{9\sqrt[3]{m}}-\frac{1}{\sqrt[3]{m^4}}\left(\frac{4}{27}C+\frac{1}{162}\right)\cdots$$

If $C=\frac{1}{-24}$ then

$$\text{Var}(Y)\cong\frac{1}{9\sqrt[3]{m}}$$

Let

$$YNR = \sqrt[3]{N-1/24} \qquad N \geq 1$$

$$Z = \frac{YNR - \sqrt[3]{m+1/24}}{\frac{1}{3\sqrt[6]{m}}}$$

$$YNR = \sqrt[3]{m+1/24} + (1/3\sqrt[6]{m})Z$$

$$\sqrt[3]{N-1/24} = \sqrt[3]{m+1/24} + (1/3\sqrt[6]{m})Z.$$

Thus set

$$N = \begin{cases} 0 & \text{if}\left(\sqrt[3]{m+1/24} + \frac{1}{3}\sqrt[6]{m}Z\right)^3 + \frac{1}{24} < 1 \\ \left[\left(\sqrt[3]{m+1/24} + \frac{1}{3}\sqrt[6]{m}Z\right)^3\right] + \frac{1}{24} & \text{otherwise} \end{cases}$$

N is now the approximated Poisson variate.

Ⅲ. Evaluation of the Methods

Generally, generating Poisson random number from the exact method is known to take a long execution time, since one generated random number "n" requires on the average n multiplications of uniform (0, 1) random numbers. Therefore, generating Poisson variables with the exact method (which gives the better accuracy) is good for a small mean, m, while the approximation methods, which take shorter execution time but with less accuracy, should be used for large m. Here we need a trade-off between execution time and accuracy to choose the generation method according to the mean value of the Poisson distribution.

The comparison statistics show how closely the methods approximate the original Poisson distribution. In the Kolmogorov-Smirnov test, all approximations are accepted at significance level $\alpha = .05$. From the comparison of the empirical probability distributions and the mean squared deviations from the exact distribution, the optimal generation procedure based on the mean value, m, is as follows:

Method	**Mean(m)**
1. exact Poisson distribution	if $0 \leq m \leq 20$
2. square root transformation	if $20 < m \leq 100$
3. normal approximation	if $m > 100$.

The cube root transformation should not be adopted because it is far less accurate compared to the square root transformation and the normal approximation. The following tables and figures show the comparison statistics of the three approximations versus the exact Poisson distribution.

Table Ⅰ assesses the accuracy of the normal, square root, and cube root approximations to the exact Poisson distribution at selected points. The mean squared deviation is defined as:

$$\left\{\frac{1}{k-1}\sum_{i=1}^{k}(P_i' - P_i)^2\right\}$$

where

P'_i is the cumulative distributed probability of the approximating distribution;

P_i is the cumulative distributed probability of the exact Poisson distribution; and

k is the sample size.

Table Ⅱ is a summary of a comparison of the empirical distributions produced by the three approximation methods to the exact Poisson distribution by means of the one sample Kolmogorov-Smirnov test. The null hypothesis is that the approximated distributions are Poisson against the alternative that they are not Poisson.

Finally, Table Ⅲ analyzes the sensitivity of the square root transformation to the constant C. In the derivation of the square root transformation, C was chosen as 3/8. Different constants were used in order to find the most robust constant to use, i.e., the constant which yielded the smallest mean squared deviation from the exact Poisson. The value $C = 13/18$ was found to be the most robust by the sensitivity analysis of mean squared deviation and poisson mean in the range of m from 20 to 100. Note that this value of C was used in the approximation when making the comparisons with the other methods.

Table I. Accuracy of the Normal, Square Root, and Cube Root Approximations to the Exact Poisson Distribution.

Poisson mean, m	Observed Value, n	$P(Z_{(m)} \leq n)$			
		Exact	Normal Approx	Square Root Transformation C=13/18	Cube Root Transformation
5	2	0.08422	0.07301	0.07719	0.10491
	3	0.14037	0.11939	0.14051	0.18491
	6	0.14622	0.16036	0.14614	0.16681
15	10	0.04861	0.04485	0.04704	0.09045
	14	0.10244	0.09937	0.10288	0.05512
	18	0.07062	0.07622	0.07073	0.06348
20	14	0.03874	0.03633	0.03746	0.08819
	19	0.08884	0.08683	0.08891	0.08726
	22	0.07692	0.08050	0.07672	0.05755
40	36	0.05394	0.05161	0.05409	0.02552
	39	0.06296	0.06223	0.06307	0.06336
	42	0.05850	0.05995	0.05841	0.05675
61	55	0.03960	0.03910	0.03963	0.00662
	59	0.05019	0.04999	0.05020	0.04973
82	82	0.04407	0.04402	0.04402	0.03289
	87	0.03686	0.03649	0.03683	0.03801
	89	0.03165	0.03243	0.03149	0.03991
90	94	0.03775	0.03805	0.03759	0.0255
	97	0.03112	0.03182	0.03098	0.0435
100	93	0.03223	0.03241	0.03217	0.0410
	99	0.03997	0.03980	0.03994	0.03605
120	112	0.02881	0.02847	0.02866	0.03627
	116	0.03477	0.03438	0.03462	0.00130

Table II. One Sample Kolmogorov-Smirnov Test.

Poisson mean, m	Sample Size	$MAX\lvert\bar{F}(x)-F(x)\rvert$			Critical Value, at $\alpha=.05$	Accept or Reject
		Normal Approx.	Square Root Transformation C=13/18	Cube Root Transformation C=1/24		
10	19	0.021	0.010	0.092	0.312	Accept
30	32	0.012	0.006	0.052	0.240	Accept
50	40	0.009	0.005	0.040	0.215	Accept
70	47	0.008	0.004	0.034	0.198	Accept
90	52	0.008	0.005	0.030	0.189	Accept

Table III. Sensitivity of the Square Root Transformation to the Constant C for Poission Distribution with Means m=20 and m=80

C	Mean Absolute Deviation $\sqrt{\frac{1}{k-1}\sum_{i=1}^{k}(P'_i-P_i)^2}$	
	Poisson Mean=20	Poisson Mean=80
3/8	0.019	0.009
5/9	0.016	0.006
7/12	0.014	0.005
11/18	0.011	0.005
23/36	0.009	0.004
8/12	0.007	0.003

*13/18	0.006	0.002
9/12	0.007	0.003
5/6	0.011	0.004

* optimal constant for C.

Acknowledgements

The author would like to express his appreciation for the many helpful comments and suggestions of Peter A.W. Lewis whose interest simulated much of this paper. Acknowledgments are also made to Donald P. Gaver and Gerald P. Learmonth for their continued interest and support.

REFERENCES

[1] Cox, D.R. and Lewis, P.A.W., *The Statistical Analysis of Series of Events*, London, Methuen, 1966.

[2] Lewis, P.A.W., Goodman, A.S. and Miller, J.M., Pseudo-Random Number Generator for the System/360, *IBM Systems Journal*, No. 2, 1969.

[3] Learmonth, G.P. and Lewis, P.A.W., *Naval Postgraduate School Random Number Generator Package LLRANDOM*, research report NPS 55 LW 7306lA, Naval Postgraduate School, Monterey, California, June 1973.

[4] Anscombe, F.J., "The Transformation of Poisson Binomial and Negative Binomial Data," *Biometrika*, Vol. 35, December 1945.

[5] Kolmogorov, A.N., *Elements of the Theory of Functions and Functional Analysis Graylock Press*; Academic Press, 1957-1961.

[5] Lancaster, H.O., *The Chi-Square Distribution* New York, Wiley, 1969.

[7] Ahrens, J.H. and Dieter, U., *Non-Uniform Random-Numbers*, to be published.

[8] Breiman, L., *Statistics with a View Toward Application*, Houghton Mifflin, 1973.

[9] Barlett, M.S., The Square Root Transformation in the Analysis of Variance, *Journal of the Royal Statistical Society, Supplement*, Vol. 3, No. 68, 1936.